Big data analytics in social care provision: spatial and temporal evidence from Birmingham

Sarunkorn Chotvijit*
Warwick Institute for the Science of Cities
Coventry, UK
S.Chotvijit@warwick.ac.uk

Malkiat Thiarai
Warwick Institute for the Science of Cities
Coventry, UK
M.S.Thiarai@warwick.ac.uk

Stephen Jarvis
Warwick Institute for the Science of Cities
Coventry, UK
S.A.Jarvis@warwick.ac.uk

ABSTRACT

There is significant national interest in tackling issues surrounding the needs of vulnerable children and adults. At the same time, UK local authorities face severe financial challenges as a result of decreasing financial settlements and increasing demands from growing urban populations. This research employs state-of-the-art data analytics and visualisation techniques to analyse six years of local government social care data for the city of Birmingham, the UK's second most populated city. This analysis identifies: (i) service cost profiles over time; (ii) geographical dimensions to service demand and delivery; (iii) patterns in the provision of services, and (iv) the extent to which data value and data protection interact. The research accesses data held by the local authority to discover patterns and insights that may assist in the understanding of service demand, support decision making and resource management, whilst protecting and safeguarding its most vulnerable citizens. The use of data in this manner could also inform the approach a local authority has to its data, its capture and use, and the potential for supporting data-led management and service improvements.

CCS CONCEPTS

• **Information systems → Geographic information systems**; **Data analytics**;

KEYWORDS

Local authority, social care, service provision, data analytics, spatio-temporal analysis, Birmingham

ACM Reference Format:
Sarunkorn Chotvijit, Malkiat Thiarai, and Stephen Jarvis. 2018. Big data analytics in social care provision: spatial and temporal evidence from Birmingham. In *dg.o '18: dg.o 2018: Proceedings of the 19th Annual International Conference on Digital Government Research, May 30-June 1, 2018, Delft, Netherlands.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3209281.3209300

*Chotvijit is the corresponding author of this research.

dg.o '18, May 30-June 1, 2018, Delft, Netherlands

ACM ISBN 978-1-4503-6526-0/18/05…$15.00
https://doi.org/10.1145/3209281.3209300

1 INTRODUCTION

Recent examples exist for supporting city governance through the collection of large, heterogeneous data sources and the application of spatio-temporal data analytics [2, 21], including in the context of budget management [31].

New technologies and techniques have been shown to influence the ability to explore and enhance communities' living conditions [2]: Research by Symons [32] for example, suggests that social care data from a local government institution could improve service delivery to individuals, and focus support to those with the most significant need. There is also evidence of the use of personal data held by the government to support service provision and planning, particularly to vulnerable children [15, 29, 30].

Issues surrounding children's social care have been widely reported by several local authorities in the UK. Serious case reviews into child deaths have led national governments to change the way in which social care performance is monitored and evaluated [4]. This has resulted in exposing inefficiency and poor allocation of safeguarding services [25].

The financial cuts to social services experienced by local authorities in England [1, 5, 13, 17, 26] and changes in the number of people receiving services [11], has resulted in difficulties in supply and eligibility [14], a reduction in service quality [19] and an incremental increase in the number of self-funded recipients [26]. The problem of identifying and examining care quality [20] has been limited by the inaccessibility of good data sources [33], insufficient high-quality care [35] and an increase in deprivation rates [6, 16].

This research has been undertaken in collaboration with Birmingham City Council, the largest local authority in Western Europe. Birmingham City Council recognises, like all local authorities, that it collects and stores a considerable amount of data on its citizens, their circumstances and the services that they consume. The intent of this research is to apply spatio-temporal data analytics to help in the understanding of social care service provision, to support resource management and budget challenges in different areas of the city, and to highlight the value in the data it obtains and stores and what can be learned from this without additional costly business analytic or consultancy services.

We note that this research does not seek to identify risk factors that affect specific individuals. Rather it aims at supporting the organisation in understanding patterns, insights and trends, which can impact future government operations to support service and budgetary planning.

2 BACKGROUND LITERATURE

In recent years, there has been growing interest in using social care data to highlight increased requirements, improve decision making processes and transforming public services [12]. This has required involvement from different stakeholders on the use of complex and personal data for search, analysis and visualisation [22]. However, using data and advanced technologies has limitations if the data quality and standards are poor [36]. This said, combining data analytics with policymaking and the design of public service within the public sector have been shown to improve service delivery [34].

Many studies have used spatial analysis, most notably Geographical Information Systems (GIS), to analyse issues related to children in social and health care. The research of Susan [9] showed biophysical and social vulnerabilities of Georgetown County using methods to calculate an index score using several indicators such as census, demographics and housing status, before using GIS mapping to create a better understanding of vulnerable areas.

In research by Ernst [10], the rates and distributions of three types of child maltreatment (physical, sexual abuse and neglect) in areas as small as neighbourhoods were identified to highlight at-risk communities. The approach adopted used least squares multiple regression analysis and GIS to gain a more detailed understanding of the child welfare system under investigation.

Child service provision plans, including for interventions and funding, were analysed by Besag and Newell [3] using a novel scoring and clustering technique. Their work supported the detection of rare events, by computing the probability of the number of observed cases given the population at risk.

Dasymetric distribution techniques have been used to spatially represent healthcare outcomes. This research highlighted significant differences between rural and urban areas [18]. While research from Yu [37] highlighted substantial respiratory health problems from child residents near Asian dust storms. The risk was specified using structured additive regression on demographic data and data including air pollutant parameters. The vulnerability of regions to increased infant mortality has previously been identified using Thiessen (Voronoi) polygons, the use of the Moran index and the G-test [28].

The use of personal information owned by the local government could create several challenges around applying comparative methods or analytics. For example, the issue of data uncertainty was introduced in Child Service Planning in Sheffield, as only half of the datasets supplied by partner organisations met the granularity requirements needed for their analysis [8]. It has also been identified that the difference in characteristics of children registered for children's social care across the UK could cause difficulties when analysing and interpreting published statistics [23].

2.1 Research Challenge

Social care is a primary and fundamental service provided by local government to support citizens with illness, disability, ageing and vulnerability within their locality. The demands on social care in the UK are increasing, while austerity policies across the UK have seen a reduction in funding and services available to meet the needs of the growing population.

We focus our research on the city of Birmingham, the UK's second largest and most populous city outside of London, with a population of over 1.1 million people [24]. Birmingham City Council (BCC) is responsible for the governance of the city, which is managed through the division of the city into 10 council constituencies and 40 electoral wards. Many issues including demographics and citizens' living conditions have increased pressure on the local authority; the council is ranked sixth in the most deprived local authorities in the UK, almost half the population are aged under 30, and 30% of children live in a deprived household [7]. Life expectancy between the most and least deprived areas in the city is 6.9 years for both sexes.

Revenue and expenditure in 2016/17 were £3.094 billion, of which more than 50% was spent on services for people and schools. Managing BCC's priorities - including maximising the independence of adults, sustaining neighbourhoods, and growing the economy and jobs - has been challenging in the context of these fiscal constraints. The council is expected to make total savings of £815 million over the 9 years period 2011/12 to 2019/20. This will lead to a reduction in the total number of staff from 20,000 in 2010 to around 7,000 by 2018.

This research was predicated on the basis that the council was seeking to better understand the demand for its services and how this demand could be met or managed during a period where financial resources were diminishing. The council was rated inadequate in service provision of social care to its citizens by the UK Office for Standards in Education, Children's Services and Skills (Ofsted) [27]. The organisation was cited nationwide for serious failures in protecting and safeguarding of vulnerable people. A number of high-profile child deaths in Birmingham have been reported in the national press since 2003.

In order to illustrate the financial pressures on the council, which compound the operational issues described, Table 1 illustrates the plan for budget reduction for Birmingham City Council compared to other local governments in England in 2016/2017.

Table 1: Budget reduction plan for the top 5 councils in England in 2016/2017

City Council	Budget Reduction Plan
Birmingham	£88.2 million
Bristol	£52 million
Sheffield	£51 million
Newcastle	£32 million
Liverpool	£28 million

BCC, like many other local authorities, has sought to make better use of the data that it holds to enhance city governance and use it in different contexts, such as in financial planning and street cleaning optimisation. It is also making aspects of the data 'open' as part of its transparency agenda.

This study employs spatial and temporal analysis which, to our knowledge, has not previously been applied to social care data in the manner intended by this research. To begin the process of accessing the data, the researchers followed the council's internal governance processes to ensure full compliance with relevant data protection and ethical obligations. The research was conducted as part of the authority's Future Council Programme, was based at BCC's offices in Birmingham and, had the support of senior leaders in the council. The research aimed to investigate:

(1) How data held in local authority systems could be exploited to provide significant value and insight to the local government and community;
(2) The extent to which data value is impacted when personally identifiable attributes are retained at the most fine-grained level of analysis;
(3) How the use of local authority data could inform future planning and service delivery in Birmingham, as part of the authority's business planning and budget setting processes.

The research used the closed agreements dataset from BCC's internal case management database called CareFirst, which has records of social care provision for all registered citizens over the past 15 years. An agreement refers to a commissioned delivery of a social care service following an assessment, according to an individual level of need and eligibility criteria. As of March 2016, the system had agreements numbering more than 600,000 client records.

3 CARE SERVICE AGREEMENTS

This research uses data extracted from the CareFirst structured closed agreements. The dataset included 31,610 unique clients, registered for 119 different council services that were further sub-divided into 360 service elements. Each closed agreement comprises 11 attributes, see Table 2.

Table 2: Records comprising a closed agreement and their description

Record	Description
ADEID	Agreement ID
PERID	Person ID
DOB	Date of Birth
Agreement Start	Start date of agreement
Agreement End	End date of agreement
Service	Alphanumeric coding of service
Service Description	Description of service
Element	Alphanumeric coding of element
Element Description	Description of element
Postcode	Postcode (unit level)
Weekly Cost	Cost per week per one agreement element

ADEID and PERID normally appear in integer form. A person ID can be duplicated and can include one or more ADEID (agreement record in the CareFirst system) attached to the individual, but not vice versa. The service and element names are typically stored as a string comprising five or more characters, representing a short version of the full description. A simple coding strategy is employed: A name that begins with CH is related to children; DIR represents a direct payment; HSSU represents home support; LD is related to learning disabilities; MH is related to mental health; OA refers to a service element for an older adult; PD represents a service for people with physical disabilities and, SM represents a service connected to substance misuse. For example, CHEFODIS stands for Children External Fostering Disabled and PDEHSUPP represents the service for Physical Disabilities External Support Living. Postcode details are used in this research but are configured to allow us to preserve the anonymity of individuals but, at the same time, be fine-grained enough to provide meaningful spatial analysis. A postcode can be divided into three levels: district, sector and unit. An example of a relevant postcode is 'B1 1AA'. The district postcode accounts for those letters and numerals before the space, representing part of the city, in this case 'B1'. While the sector code includes one more numeral after the space to display a deeper level sub-area of that district, in this case 'B1 1'.

The data sample used in this study represented expenditure by the Council on service agreements totalling over £100 million.

4 METHODOLOGY

This study uses spatio-temporal analysis, which can be applied when data are collected in time (interval) as well as space (location and geometry), in our case this includes start and end dates of service agreements and postcodes. The analysis excluded any possible outliers from the data, which could otherwise distort the overall results. The research was structured using a number of different methods. Firstly, we focussed on the top twenty service elements, having determined through analysis they accounted for approximately 80% of the total cost of all service elements. The annual cost analysis of the top twenty elements highlighted that the cost of these services had fallen by about £1 million over the study period. Given the Council's budget reduction plan, this trend was likely to continue. Secondly, the data with postcodes were analysed to highlight geographical areas across the city with the highest cost and demand, and a matrix heat map was used to identify hot spots for further analysis. Thirdly, spatial density heat maps were then generated for different age groups and services, which themselves can be animated to show the demand for services in the city over time. Finally, the geographical demand was assessed, to help understand those services where historical demand has been high by location and what conclusions can be drawn from this to inform future service provision.

4.1 Data cleansing and data pre-processing

The matrix heat map highlighted data quality and allowed us to conduct anomaly detection. A heat map in Figure 1 displays the frequency of registered agreements for all registered recipients over 75 (of the 79) postcode districts. Colour gradation highlights a higher number of agreements within a postcode district over a one year period. The remainder of this paper uses the most recent 6 years of CareFirst data from March 2010 onwards. This approach and the resulting heat map enabled the researchers to identify areas of the city where provision of services were concentrated, to focus more detailed analysis of the data.

Our data were further categorised into four age groups according to council norms: Children aged 0-11; Young People and Adults, aged 11-25; Adults aged 25-65 and Older People aged 65-90. Records are retrieved for these age ranges, see Table 3; note that there are some duplications of individuals, as a person may be registered for more than one service within a year.

Table 3: Number of service agreement records by age groups

Age Groups	Number of Agreement Records
0-11	7,308
11-25	26,142
25-65	47,247
65-90	133,599

Postcode	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
B1	0	20	27	38	23	25	36	42	46	42	43	69	61	55	42	0
B10	0	69	141	152	145	257	326	407	387	336	302	341	402	444	276	0
B11	1	96	219	266	178	273	365	487	500	530	443	556	492	477	353	1
B12	0	144	127	249	154	162	188	310	276	261	255	352	267	330	237	0
B13	1	421	305	622	414	413	653	803	820	698	623	697	644	623	505	0
B14	0	275	314	545	609	734	909	1087	1016	952	833	893	885	862	720	0
B15	0	65	64	107	101	140	172	179	154	159	125	190	154	100	113	0
B16	1	104	126	204	159	167	213	254	287	230	249	246	267	251	199	0
B17	5	192	163	236	272	304	428	455	400	462	445	511	479	421	311	1
B18	1	127	120	203	130	177	180	167	211	332	234	208	240	235	172	0
B19	1	67	167	164	134	184	187	273	315	290	298	318	397	444	400	0
B2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
B20	1	302	231	420	177	332	406	659	508	527	432	500	535	563	435	0
B21	0	118	176	199	130	261	372	517	451	419	423	417	383	457	315	1
B23	4	292	259	491	424	757	1007	1095	1056	926	924	1073	891	849	556	1
B24	4	398	324	587	297	632	637	741	772	690	668	727	739	598	380	0
B25	0	58	172	163	110	208	263	262	256	256	235	248	281	300	184	1
B26	0	122	186	210	181	422	622	611	753	659	625	597	504	599	461	0
B27	1	207	227	333	265	327	451	542	508	524	449	559	470	488	309	1
B28	0	117	146	202	289	319	507	574	537	488	433	570	451	461	381	2
B29	0	257	285	493	506	566	668	728	743	607	468	663	594	493	335	0
B3	0	0	0	1	1	0	0	4	0	0	5	3	5	7	8	0
B30	0	284	239	471	375	447	569	584	639	614	571	653	587	556	396	0
B31	0	216	341	456	662	821	1029	1066	1196	1098	1046	1028	1076	999	740	2
B32	0	125	246	293	395	580	692	665	732	718	634	847	798	781	592	2
B33	4	328	389	594	362	655	789	760	824	767	623	669	649	654	434	1
B34	2	103	148	187	137	409	466	498	705	537	482	495	465	450	373	0
B35	1	49	104	118	117	236	236	311	299	271	260	311	246	237	180	0
B36	1	77	99	92	60	133	205	192	219	206	217	208	231	213	114	0
B37	0	26	27	39	20	16	4	7	23	17	16	12	22	14	3	0
B38	0	217	243	380	361	415	419	538	542	511	491	512	436	401	314	0
B4	0	2	2	9	1	2	3	2	5	6	4	6	3	3	1	0
B42	2	56	137	146	204	253	405	430	387	403	366	375	389	334	289	0
B43	0	3	7	13	12	11	10	8	16	22	21	33	21	8	0	0
B44	5	224	336	528	685	701	907	857	794	750	693	771	672	633	504	0
B45	1	107	182	205	210	294	390	334	452	373	345	337	314	340	297	1
B46	0	10	5	15	10	5	4	2	7	6	7	9	10	3	4	0
B47	0	5	0	3	2	1	5	4	3	7	5	4	3	1	1	0
B48	0	24	4	29	10	10	18	23	21	5	11	9	9	4	2	0
B49	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
B5	0	16	49	51	85	108	203	169	207	195	177	160	208	155	145	0
B50	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0
B6	2	66	80	103	47	120	169	190	274	253	243	379	321	252	192	1
B60	0	6	2	12	1	7	3	6	12	6	1	4	5	1	3	0
B61	0	18	12	33	10	5	12	19	15	17	8	6	2	3	1	0
B62	0	11	8	23	14	11	31	14	18	11	10	14	11	5	9	0
B63	0	1	0	4	0	1	0	1	0	3	0	0	0	0	0	0
B64	0	3	1	5	4	1	3	1	3	0	1	1	3	0	0	0
B65	0	1	2	4	2	3	0	2	10	1	0	4	4	1	1	0
B66	0	42	48	87	21	20	47	50	35	9	32	44	15	18	7	0
B67	0	2	7	19	3	2	9	11	5	8	5	9	10	3	1	0
B68	0	12	2	13	2	2	1	8	9	9	8	11	3	0	0	0
B69	0	13	17	19	9	10	11	7	16	6	0	8	3	1	1	0
B7	0	56	67	98	74	102	123	118	154	148	127	169	156	146	113	0
B70	0	20	24	32	18	4	7	18	21	18	12	30	16	5	5	0
B71	0	11	5	19	15	6	5	12	17	7	4	4	4	1	2	0
B72	0	83	53	122	119	131	165	211	180	177	159	132	117	109	103	0
B73	1	150	199	331	347	418	509	466	477	386	342	353	286	311	207	0
B74	0	134	113	207	254	267	316	411	427	326	288	232	198	197	160	0
B75	2	88	242	282	391	452	629	480	602	561	468	456	356	319	262	0
B76	0	56	151	223	318	343	418	420	460	424	355	313	318	300	229	0
B77	0	6	7	14	10	6	4	6	5	14	9	6	7	8	3	0
B78	0	5	2	7	3	5	1	6	0	0	0	3	4	0	0	0
B79	0	18	9	32	7	16	14	7	9	9	6	5	4	2	0	0
B8	3	146	142	223	133	292	404	476	544	454	335	399	387	455	409	1
B80	0	0	3	2	0	0	0	1	0	0	0	0	0	0	0	0
B9	5	111	146	175	141	202	216	303	299	289	201	316	315	342	272	1
B90	0	24	23	29	17	10	1	2	10	13	10	8	4	6	3	0
B91	0	15	1	7	6	1	3	2	59	1	5	4	4	2	0	0
B92	0	60	28	70	44	29	31	26	35	15	17	17	23	10	8	0
B93	0	6	1	4	0	1	0	2	0	1	2	1	0	1	0	0
B94	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
B95	0	2	2	4	1	4	1	2	2	2	0	0	0	0	0	0
B97	0	33	11	35	12	12	7	3	16	3	8	7	9	3	3	0
B98	0	1	0	0	1	0	1	4	1	1	2	0	0	0	0	0

Figure 1: Matrix frequency heat map of all service records in all postcode districts over the past 15 years

We illustrate our data-processing workflow that consists of data pre-processing, trend analysis and geographical mapping as follows:

(1) Data ingestion, cleansing and anomaly detection: We used R scripts to remove erroneous characters, conduct range checks and identify missing values. Of the 258,673 closed agreements studied, 18,872 (7.3%) were removed because of 'bad data': The process mostly involved missing values, unreadable or invalid data records, unknown or invalid age and gender, and service users from non-Birmingham postcodes.
(2) Trend analysis: We also used R scripts to visualise interesting patterns found in the data. This included population, the area of interest and cost of service element analysis.
(3) Geographical mapping: We employed the open-source geographical information systems QGIS to perform spatial-temporal mapping using postcodes in the closed agreements.

As the data contained postcodes, exploration was possible at the sector level, at which point the data was spatially joined with a geographic shapefile (in the ESRI vector data storage format) representing the location, shape and attributes of the corresponding geographic unit. Coordinates were plotted using the Ordnance Survey National Grid reference system (BNG) with the European Petroleum Survey Group (EPSG) Code EPSG:27700. Plugins for OpenStreetMap was employed from the QGIS OpenLayers Plugin. [1]

5 IDENTIFICATION OF THREE REGIONAL HOT SPOTS

In order to better understand the provision of services and their cost, the focus of the analysis was conducted at the district level to retain privacy of the individuals. Two methods were required to choose the candidate areas, so that (i) the areas showed the highest density of recipients across all age groups and (ii) areas had a population of approximately 50,000 people.

Figure 2 displays the point and density heat map of the entire social care client records in the extracted dataset. Each density heat map is plotted by the 4 different age groups (columns) and represents the service provision over two three-year periods (rows) to enable comparison over time: Dark color shows occurrence and represents density. This helped identify potential locations with high demand (darker spots indicate higher density of service provisions). For example, the service provision for children aged 0-11 (the first column), even though number of agreement records is far fewer than other age bands, the provision tends to cluster at some parts of the city and those differ over the two three-year periods. Whereas, the maps representing the 11-25 age group (the second column) demonstrate a good dispersal of service agreements over the periods as the services distribute quite evenly throughout the city (fewer hot spots), though some are concentrated around the southern region. However, for registered individuals aged 25 and over (the last two columns), there is a greater diversity in clients receiving different service elements from the authority.

Consequently, a second method was applied to help identify the areas of interest that may contain multiple district-level postcodes. Estimating the population per postcode was calculated by aggregating the number of citizens within a particular district from the 2011 Census of Postcode Headcounts and Household Estimates, before retrieving an approximate number of people per postcode. Regions that displayed the highest density of registered service users and where the residents totalled approximately 50,000 were located in the eastern, northern and southern parts of central Birmingham, see Figure 3. In addition, Table 4 shows the postcodes involved and the approximate aggregated population from each chosen region. Note that, the final decision when choosing the candidate area also considered the number of adjacent postcodes as well as the size of the corresponding focussed region.

We produced further analysis of the three regions by examining the annual cost of the top twenty service elements in all age groups over the six-year period.

[1] Polygons of the UK boundaries can be obtained from the UK Data Service.

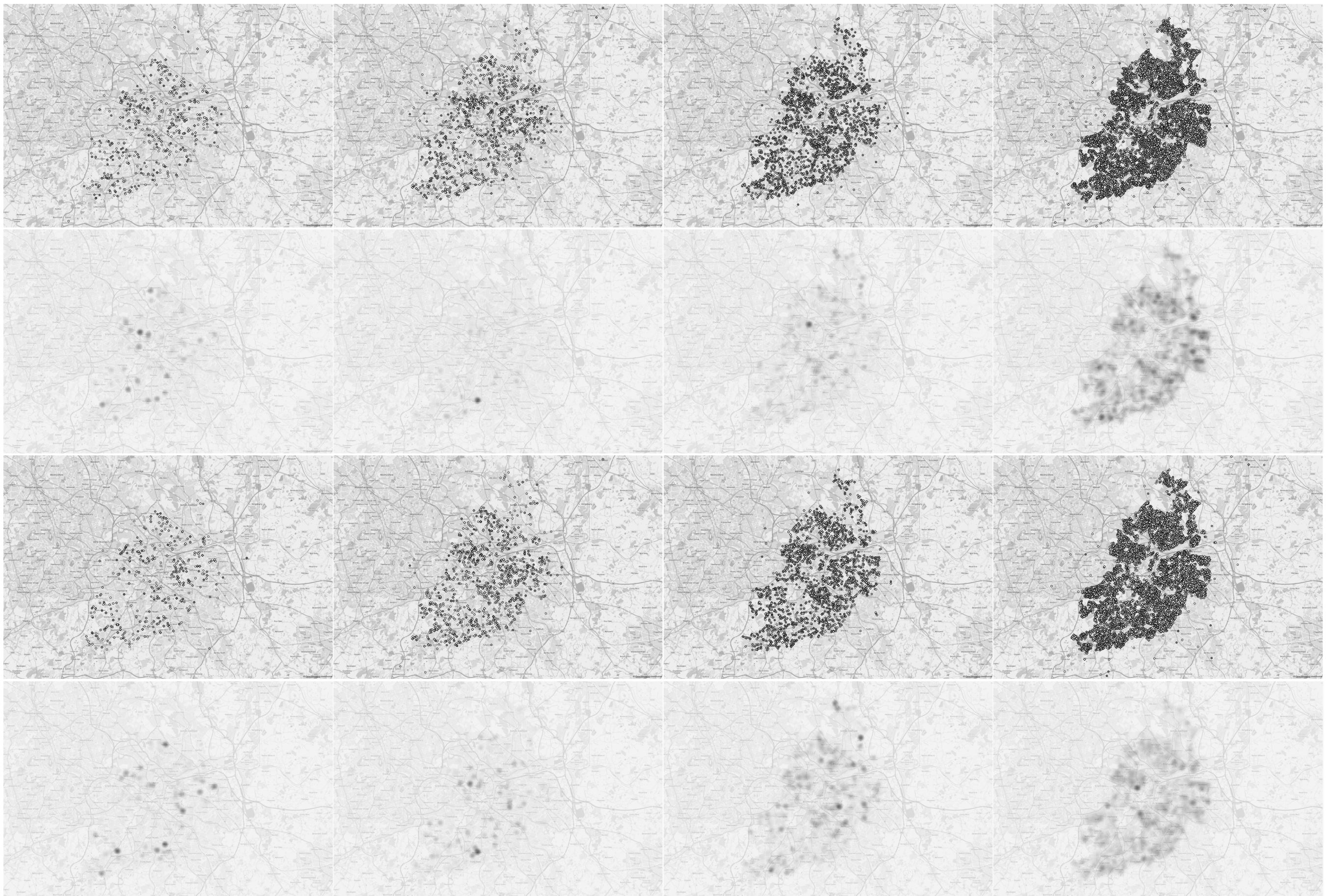

Figure 2: The location of social care service provision across Birmingham, ages 0-11 (1st column), 11-25 (2nd column), 25-65 (3rd column) and 65-90 (4th column) from 2010-2012 (first two rows) and 2013-2015 (last two rows) both point and density heat maps

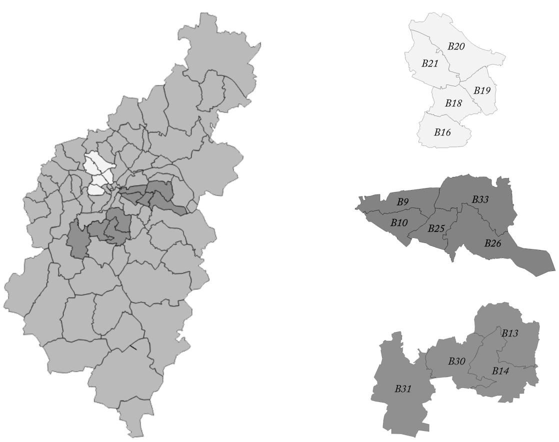

Figure 3: Birmingham postcodes at a district level, with 3 areas of interest; northern, eastern and southern

Table 4: Postcodes within the 3 areas of interest and the estimated population

Regions	Postcodes	Population (approx.)
Northern	B16, B18, B19, B20, B21	45,000
Eastern	B9, B10, B25, B26, B33	45,000
Southern	B13, B14, B30, B31	65,000

5.1 Northern

The northern part of the city is smaller by size but was included as it represents a higher density of population per district. Figure 4 shows the annual cost of service provision for the top 20 social care elements for the northern region. Other than a small decrease in cost in 2012, the overall spending by the local government increases by almost 30% between 2010 and 2015. Moreover, the pattern emphasises registered clients aged 11-25, which after 2010 grow by more than 120% to the end of 2015.

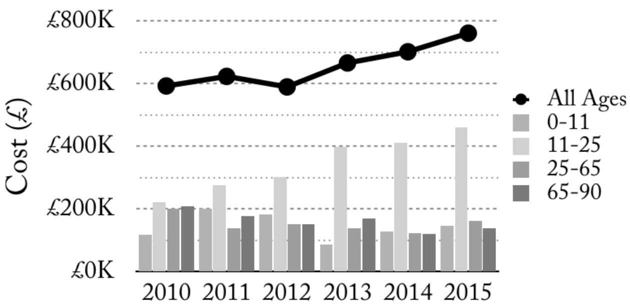

Figure 4: Cost of top 20 elements for the northern area

5.2 Eastern

In the eastern part of the city, service element provision is dispersed evenly across the region over time. Figure 5 indicates an increasing pattern of overall cost starting from an initial increase in 2011 to 2012 of approximately 24% before gradually diminished to less than £750,000 in 2015. The 11-25 age group, as in the northern region, dominates from the start of the period (with costs in excess of £250,000) and continues to increase to more than 50% of the overall cost by 2015.

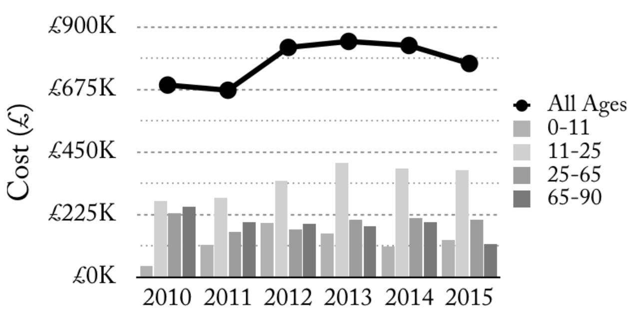

Figure 5: Cost of top 20 elements for the eastern area

5.3 Southern

In the southern part of the city, fewer postcodes were selected for analysis. However, the region contained more unique recipients of services as well as greater overall population. The cost of service element provision, found in Figure 6, is therefore higher than the other two regions. The data shows a slow reduction in the accumulated cost from above £1.4 million at the beginning of the period to around £1.35 million by the end of the period. Despite a decreasing trend in the annual cost for the top 20 service elements, the cost for clients who are aged 11-25 has increased by approximately 30% from 2010 and dominates other age groups.

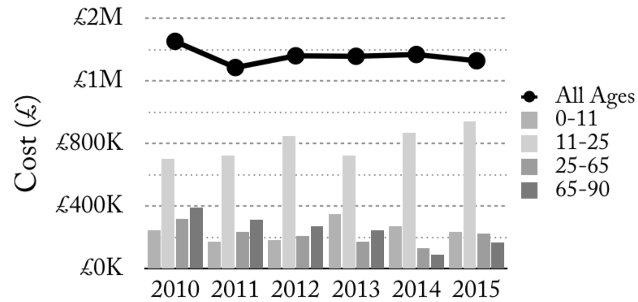

Figure 6: Cost of top 20 elements for the southern area

The cost profile for other age groups shows fluctuating yet diminishing costs over the six-year period. The total cost of services for each region over the six years period is: northern region £3.9 million; eastern region £4.6 million and southern region 8.1 million.

Our analysis subsequently investigated the geographical dispersal of the top 20% of recipients who receive the largest number of service agreements over two three-year periods, see Figure 7. A threshold percentage was used to highlight the top quartile of data and techniques were employed to ensure individual privacy. The data was overlaid over the district postcode boundaries of the three regions as shown in Figure 3. The height and colour intensity of each cylindrical bar are determined by the number of unique agreements registered within the postcode region; the higher and darker bar indicates the higher number of agreements involved. The change of service provision between two periods is apparent.

The use of the 3-D map has helped us identify not only the location of service users, but also the frequency of services within those regions and, in particular, the temporal demand.

6 DISCUSSION

This research seeks to identify new ways of integrating data analysis and visualisation with the monitoring and analysis of social care service closed agreements. This work is being used to support Birmingham City Council in resource and service allocation, a task which is particularly challenging in times of financial austerity. This data is not 'open data', but is stored in BCC's data archives and there was a desire to make use of this data to better understand service provision at different levels of granularity, develop new analysis to predict future demand and, improve the strength of data interpretation to identify risk.

The research presented here focusses on applying spatio-temporal analysis to three geographical areas of interest, which provide an overall picture of where council spending on these services has taken place, and the age groups of registered users. We can identify the distribution of services over time and, in our example, the uptake of these services by people who are aged 11 to 25 across the city. Child care services (for those aged below 18) are notable, as they dominate all other groups in the cost of social care services.

The analysis of the three city regions demonstrates that in the southern part of the city, service payment per year was approximately 25% more than in the northern and eastern parts of Birmingham. Despite a higher number of agreement records, services

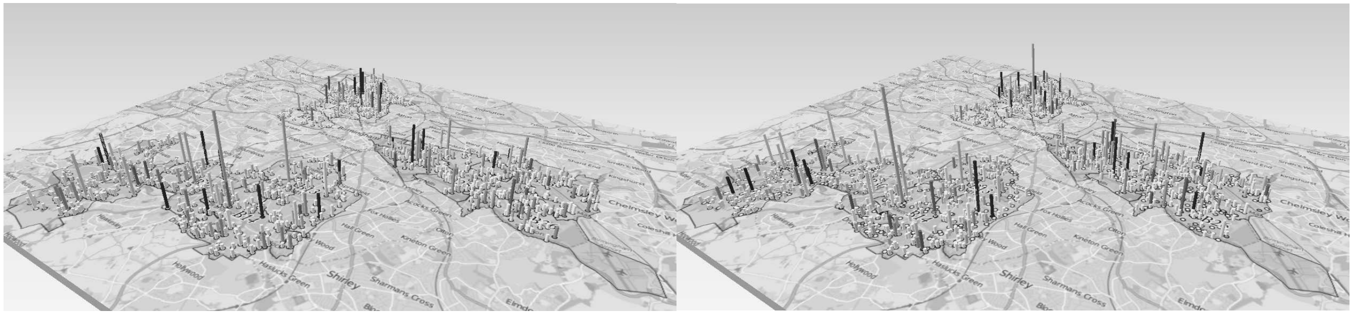

Figure 7: 3-D geographical dispersal of service agreements for all age groups across three regions, 2010-2012 (left) and 2013-2015 (right)

for older adults (aged over 25) were less costly than those of the younger age group. The cost patterns and three-dimensional dispersal map emphasise the impact of clients aged 11-25 in the number of recipients and agreement records received from different locations across the three city regions. We found that this age band also shows that council spending is highly correlated to the region's total service provisioning cost.

These results provide detailed analysis of the financial commitment that the local authority is required to make in specific regions in the city. Despite the overall annual spend declining over the past six-years, support within the northern region has increased by 33%. These remain a topic of further investigation by the council. In presenting the service provisioning at the district postcode, a more fine-grained analysis is possible, while avoiding disclosure of recipient's identity and data. It is possible to re-apply the techniques used in this research and repeat the analysis with different parameters, for example, analysing other age ranges (0-5, 6-11, 11-16, etc.), changing the population size or identifying the most costly services, in order to provide an alternative view of the data for the council to consider.

7 LESSONS AND CHALLENGES

We recognise that this research has presented some very interesting outcomes. The researchers acknowledge the complexity of working with social care data, as the data requires careful analysis and manipulation when applying data cleansing and pre-processing. Using new approaches to data analysis and visualisation can support the authority in exploring the service provision patterns from various perspectives by focussing on only the high demand areas. The studies show variations service demand from different age bands and the influence of this to cost in particular regions of the city. This may help the authority better target specific groups of recipients and better manage its resources for future service delivery. In addition, this research is generalisable to other city councils, in that the approach can be replicated if other authorities' systems use similar data collection and methods to classify attributes.

However, limitations still remain concerning data quality, sample size and academic resources. The number of attributes extracted from the council's systems is also somewhat limited. This is in part due to the council's need to retain personal information, as well as consideration for data privacy and individual rights. The researchers also discovered that there was also a need to include extra agreement records (extending the analysis to involve data from previous years) to better understand the historic service provisioning in the city. Although there is evidence of the use of data analytics to analyse issues related to children in social and health care, the amount of related literature that has used spatio-temporal analysis (or a similar approach) to support local government decision making and resource management within the social care service domain is limited.

8 CONCLUSION AND FUTURE WORK

UK local authorities face severe financial challenges as a result of decreasing financial settlements and increasing demands from growing urban populations. At the same time, there is significant national interest in tackling issues surrounding the needs of vulnerable children and adults. This research aims to support this challenge, in particular by using state-of-the-art data analytics and visualisation techniques to analyse six years of local government social care data for the city of Birmingham. The analysis shows: (i) service cost profiles over time; (ii) geographical dimensions to service demand and delivery; (iii) patterns in the provision of services, and (iv) the fact that significant data value can be extracted from closed data with the right data cleansing and privacy filters. Patterns and insights are presented that may assist in the understanding of service demand, supporting decision making and resource management, whilst protecting and safeguarding the city's most vulnerable citizens.

We show that the quality of data collected by the council is significant (only 7% of data is removed because of problems with the data records) and, whilst the amount of data studied here is relatively small, we highlight that a relatively small amount of data can, if extracted and organised effectively, produce results that can support a wide-ranging set of objectives.

There are some areas of further work that can be taken forward as a result of this study. These include expanding the number of small regions of analysis to gain a better understanding of service delivery at the local level; applying statistical analysis such as correlation, regression, classification or other approaches to support prediction models; identifying other datasets that complement the existing data and how this may assist and inform this debate; and using other datasets from the same or similar systems, such as detailed assessment or referral, which may provide information on *why* a service was taken up and the progress of recipients through the referral, assessment and service-delivery system.

ACKNOWLEDGMENTS

The lead author gratefully acknowledges funding support by the UK Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Urban Science and Progress under Grant No.: EP/L016400/1. The authors also thank Birmingham City Council for providing support and data access through their internship programme.

REFERENCES

[1] ADASS. 2016. *ADASS Budget Survey 2016 Report.* Technical Report. London, UK.
[2] Local Government Association. 2014. Transforming local public services: using technology and digital tools and approaches. (June 2014). Retrieved April 20, 2017 from https://www.local.gov.uk/sites/default/files/documents/transforming-public-servi-2a5.pdf
[3] Julian Besag and James Newell. 1990. The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 154, 1 (June 1990), 143–155.
[4] Marian Brandon, Sue Bailey, Pippa Belderson, Peter Sidebotham, Carol Hawley, Catherine Ellis, and Matthew Megson. 2013. *New Learning from Serious Case Reviews: a two year report for 2009-2011.* Technical Report. London, UK.
[5] Tania Burchardt, Polina Obolenskya, and Polly Vizard (Eds.). 2016. *Social Policy in a Cold Climate* (eds. ed.). Policy Press, Bristol, Chapter 9, 187–214. https://doi.org/10.1007/3-540-09237-4
[6] Paul Bywaters, Geraldine Brady, Lisa Bunting, Brigid Daniel, Brid Featherstone, Chantel Jones, Kate Morris, Jonathan Scourfield, Tim Sparks, and Calum Webb. 2017. Inequalities in English Child Protection Practice under Austerity: A Universal Challenge? *Child & Family Social Work* (July 2017), 1–9. https://doi.org/10.1111/cfs.12383
[7] Birmingham City Council. 2016. *Business Plan and Budget 2016+.* Technical Report. Birmingham, UK.
[8] Massimo Craglia, Robert Haining, and Paola Signoretta. 2003. Identifying Areas of Multiple Social Need: A Case Study in the Preparation of Children Services Plans. *Environment and Planning C: Politics and Space* 21, 2 (April 2003), 259–276. https://doi.org/10.1068/c0229
[9] Susan L. Cutter, Jerry T. Mitchell, and Michael S. Scott. 2000. Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South CarolinaRevealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographer* 90, 4 (Dec. 2000), 713–737. https://doi.org/10.1111/0004-5608.00219
[10] Joy Swanson Ernst. 2000. Mapping Child Maltreatment: Looking at Neighborhoods in a Suburban County. *Child Welfare* 79, 5 (Oct. 2000), 555–572.
[11] Jose-Luis Fernandez, Tom Snell, and Gerald Wistow. 2013. *Changes in the Patterns of Social Care Provision in England: 2005/6 to 2012/13.* Technical Report. London, UK.
[12] Michele Foster, Jennifer Harris, Karen Jackson, and Caroline Glendinning. 2008. Practitioners' Documentation of Assessment and Care Planning in Social Care: The Opportunities for Organizational Learning. *The British Journal of Social Work* 38, 3 (April 2008), 546–560. https://doi.org/10.1093/bjsw/bcl366
[13] The Kings Fund. 2015. *The Budget: Health and Social Care Funding.* Technical Report. London, UK.
[14] Caroline Glendinning. 2012. Home Care in England: Markets in the Context of Under-Funding. *Health & Social Care in the Community* 20, 3 (Feb. 2012), 292–299. https://doi.org/10.1111/j.1365-2524.2012.01059.x
[15] Jonathan Godwin. 2013. Better Data will Help Early Intervention Teams Orotect Vulnerable Children. (Dec. 2013). Retrieved December 12, 2016 from https://www.theguardian.com/social-care-network/2013/dec/05/vulnerable-children-early-intervention-data
[16] Rick Hood, Allie Goldacre, Robert Grant, and Ray Jones. 2016. Exploring Demand and Provision in English Child Protection Services. *The British Journal of Social Work* 46, 4 (June 2016), 923–941. https://doi.org/10.1093/bjsw/bcw044
[17] HSCIC. 2015. *Personal Social Services: Expenditure and Unit Costs, England - 2014-15, Final release.* Technical Report. Leeds, UK.
[18] Mitchel Langford and Gary Higgs. 2006. Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population. *The Professional Geographer* 58, 3 (May 2006), 294–306. https://doi.org/10.1111/j.1467-9272.2006.00569.x
[19] Jane Lewis and Anne West. 2017. Early Childhood Education and Care in England under Austerity: Continuity or Change in Political Ideas, Policy Goals, Availability, Affordability and Quality in a Childcare Market? *Journal of Social Policy* 46, 2 (April 2017), 331–348. https://doi.org/10.1017/S0047279416000647
[20] Juliette Malley and José-Luis Fernández. 2010. Measuring Quality in Social Care Services: Theory and Practice. *Annals of Public and Cooperative Economics* 81, 4 (Nov. 2010), 559–582. https://doi.org/10.1111/j.1467-8292.2010.00422.x
[21] Fola Malomo and Vania Sena. 2017. Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector. *Policy & Internet* 9, 1 (March 2017), 21. https://doi.org/10.1002/poi3.141
[22] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela H. Byers. 2011. *Big data: The Next Frontier for Innovation, Competition, and Productivity.* Technical Report. USA.
[23] Emily R. Munro, Rebecca Brown, and Esmeranda Manful. 2011. *Safeguarding Children Statistics: The Availability and Comparability of Data in the UK.* Technical Report. London, UK.
[24] Andrew Nash. 2015. *Statistical Bulletin: National Population Projections: 2014-based Statistical Bulletin.* Technical Report. London, UK.
[25] Ofsted. 2013. *Ofsted: Rasing standards improving lives 2012-2013.* Technical Report. Manchester, UK.
[26] Ofsted. 2014. *The impact of funding reductions on local authorities.* Technical Report. London, UK.
[27] Ofsted. 2016. *Birmingham Re-inspection of Services for Children in Need of Help and Protection, Children Looked After and Care Leavers.* Technical Report. Manchester, UK.
[28] Mirella Rodrigues, Cristine Bonfim, José Luiz Portuga, Idê Gomes Dantas Gurge, and Zulma Medeiros. 2013. Using Spatial Analysis to Identify Areas Vulnerable to Infant Mortality. *Rev Panam Salud Publica* 34, 1 (July 2013), 36–40.
[29] Laura Santhanam. 2016. Can Big Data Save These Children? (March 2016). Retrieved June 4, 2017 from https://www.pbs.org/newshour/updates/can-big-data-save-these-children/
[30] Holden Slattery. 2015. Big Data Wave Breaks on Child Protective Services. (April 2015). Retrieved February 2, 2017 from https://chronicleofsocialchange.org/los-angeles/big-data-wave-breaks-on-u-s-child-protection/10822
[31] Tom Symons. 2016. Datavores of Local Government: Using data to make services more personalised, effective and efficient. (July 2016). Retrieved January 15, 2017 from https://www.nesta.org.uk/sites/default/files/local_datavores_discussion_paper-july-2016.pdf
[32] Tom Symons. 2016. WISE COUNCIL: Insights from the Cutting Edge of Data-driven Local Government. (Nov. 2016). Retrieved February 1, 2017 from https://www.nesta.org.uk/sites/default/files/wise_council.pdf
[33] Agnes Turnpenny and Julie Beadle-Brown. 2015. Use of Quality Information in Decision-Making about Health and Social Care Services âĂŞ A Systematic Review. *Health and Social Care in the Community* 23, 4 (July 2015), 349–361. https://doi.org/10.1111/hsc.12133
[34] Barbara Ubaldi. 2013. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. *OECD Working Papers on Public Governance* 22 (May 2013), 1–60. https://doi.org/10.1787/5k46bj4f03s7-en
[35] Ivana L. Valle, Lisa Holmes, Chloe Gill, Rebecca Brown, Di Hart, and Matt Barnard. 2016. *Improving Children's Social Care Services: Results of a Feasibility Study.* Technical Report. London, UK.
[36] Chris Yiu. 2012. *The Big Data Opportunity: Making Government Faster, Smarter and more Personal.* Technical Report. London, UK.
[37] Hwa-Lung Yu, Chiang-Hsing Yang, and Lung-Chang Chien. 2013. Spatial Vulnerability under Extreme Events: A case of Asian Dust Storm's Effects on Children's Respiratory Health. *Environment International* 54 (Feb. 2013), 35–44. https://doi.org/10.1016/j.envint.2013.01.004